DOCUMENT RESUME

ED 217 078 TM 820 336

AUTHOR Schubert, Jane G.
TITLE Evaluability Assessment: The Promise in Practice.
PUB DATE Mar 82
NOTE 10p.; Paper presented at the Annual Meeting of the American Educational Research Association (66th, New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Data Collection; Evaluation Methods; Evaluation Needs; Evaluators; *Program Evaluation
IDENTIFIERS American Institute for Research; *Evaluability Assessment

ABSTRACT

Studies were conducted by the American Institute for Research (AIR) under contract with the Department of Education (ED). An evaluability assessment determines the extent to which a program is ready for evaluation, the changes needed to make the program more manageable and accountable, and toward what questions a more extensive evaluation might usefully be directed. The evaluators must learn about a specific program before performing a responsible evaluation. However, at the same time, the program staff should be educated about evaluability assessment and methodology. If this is neglected, potential benefits from participating in the study are lost. Input from project managers is necessary to insure that programs are accurately portrayed. An evaluability assessment is an evolving concept which requires informed and responsible flexibility. The extent to which consensus exists on program objectives, major program activities, or appropriate measures of performance must be determined. Evaluators interview individuals with different perspectives on the program. The analysis of these interviews provides important information. On site data collection includes formal information, as well as informal data gained through anecdotes and impressions. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED217078

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. G. Schubert

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EVALUABILITY ASSESSMENT: THE PROMISE IN PRACTICE

Jane G. Schubert

American Institutes for Research
Palo Alto, California

Paper prepared for the American Educational Research Association.

New York City

March 1982

TM 820 336

EVALUABILITY ASSESSMENT: THE PROMISE IN PRACTICE

In 1979, Joseph Wholey published Evaluation: Promise and Performance. The book was one culmination of a decade of work by a large and talented group of people at the Urban Institute. Prior to publication of this book, the group had issued a series of reports that whittled away at the dominant position in program evaluation: a position characterized by the terms "experimental design" and "impact assessment." By 1979, it appeared that large scale evaluations were necessarily reaching the verdict of "no effect."

The Wholey book provided a constructive alternative termed "sequential purchase of information." As a concept "sequential purchase of information" seemed so obviously correct that members of the evaluation community were filled with thoughts of "why had they not thought of that."

Today, the idea remains compelling; it is exactly what we ought to do. Within this framework, evaluability assessment is the essential first step. If it is not handled correctly, the foundation crumbles. This paper presents some of the real world ways that endanger this foundation.

Cronbach, et al, in their recent publication, Toward Reform of Program Evaluation, suggest that "the best way to bring about change in thought is to provoke argument."[1] They present 95 theses as a basis for such argument. One thesis (#93) states that

> the evaluator is an educator; his success is to be judged by what others learn.[2]

In the text, Cronbach amplifies this point by suggesting that the success of an evaluation ought to be judged by how it develops the learner's potential. By enlightening the learner about a particular situation or problem, more rational and wiser choices are more apt to be made. This thesis is a valuable suggestion for those engaged in program evaluation. In this paper, I build on this thesis and add a related perspective--that the evaluator's success be judged also by what the evaluator learns from

the experience that can be translated into more responsible and credible future program evaluations.

This paper raises some issues that emerge from our conduct of six evaluability assessments; staff members in five completed studies have reviewed and commented on these issues and provided examples. At least one issue emerged from each completed study. Responsibility for the compilation and discussion rests solely with the author.

All studies were conducted by the American Institutes for Research under a single contract with the Department of Education; a management and evaluation office within ED negotiated the scope of work for and monitored the progress of each assessment. The assessments examined programs managed by other offices in the Department such as Adult Education, Special Programs for Disadvantaged Students; participation in an assessment was by choice of a program manager.

It may be useful to very briefly summarize the methodology of an evaluability assessment before presenting some ideas emerging from our use of the methodology. The purpose of an evaluability assessment is to determine the extent to which a program is ready for evaluation, what changes may be needed to make the program more manageable and accountable, and toward what questions a more extensive evaluation of the program might most usefully be directed. In general, the evaluability assessment explores and documents:

- the program objectives, expectations, and causal assumptions of policymakers and managers in charge of the program;
- what political groups the program objectives as seen by others, such as (Congress, Executive Branch policymakers, and interest groups) say the program objectives are;
- the extent to which management's program objectives and information needs have been defined in manageable terms;
- the program activities actually under way;
- the likelihood that program activities will achieve measurable progress toward program objectives;

- likely uses of information on program performance; and
- options for changes in program activities, objectives, and use of information that could enhance program performance.[3]

Throughout each of the studies, interaction among staff from the sponsoring office, staff representing the program being examined, and AIR staff responsible for conducting the assessment occurred through monthly meetings of a Work Group. Members typically represented the program office, sponsor, evaluation staff, OMB, and others specifically designated. During these meetings, AIR reported on progress made, outlined plans for the ensuing month, and presented drafts of products for review and comment by all members of the Work Group. The Work Group mechanism is designed to keep all key actors informed and involved in order to enhance the usefulness of the products.

We now turn to issues arising from the experience of conducting these evaluability assessments; we view these as topics for consideration in conducting similar studies in the future.

1. Educating the Participants

At the outset of any evaluation, two kinds of information imbalance exist. The evaluators know more about the assessment methodology than the program staff and the program people are clearly more informed about what the program does and why it exists than do the evaluators. Everyone recognizes that the evaluators must learn more about a specific program before they can perform their tasks responsibly. This learning is accomplished by reading program descriptions and reports, and by talking to people who are knowledgeable about and represent various aspects of the program. In our evaluability assessment, the Work Group offered a forum for educating the evaluator, in addition to its other functions such as reviewing interim products, identifying key people to be interviewed.

Far less attention is given to educating program staff about evaluability assessment: what it can and cannot do, what options are available, and so forth. The evaluator absorbs information about the program

from the managers and staff, but devotes little attention to educating the program people about the methodology. Such educational activities are desirable in any evaluation procedure; in an evaluability assessment we believe it to be critical. Failure to educate leads to lost opportunities for program managers to know what to ask for; the consequence of this is that potential benefits from participating in the study are lost or diluted. For example, during the Work Group review of a product such as a function model, how the task fits into the overall assessment, must be clearly explained. If it isn't, these exercises seem fragmented and isolated; additional requests for similar input may also be met with little enthusiasm if the value of the investment is not understood.

2. Identifying the Appropriate Client

The institutional arrangement for the evaluability assessments being discussed were described earlier. If the program offices awarded the contracts to AIR directly, individual scopes of work would no doubt reflect a closer match to program manager requirements, along with the accompanying responsibilities of project monitoring. We recognize that such arrangements have raised questions that focus on evaluator-program relationships influencing the assessment process. Since AIR has long advocated an interactive mode for all evaluations, this pseudo-problem will not be discussed in this paper.

The point here is that the contractual arrangement, coupled with the keen interest of the sponsoring agency to develop uniformity in the process and the products may have interfered with responding to program needs. Under this contract the sponsor was supporting studies of ED programs as a service. In order for the sponsor to gain credibility within the bureaucracy, it strongly favored commonality among the products; producing this commonality consumed energy that might otherwise have been directed toward program manager needs.

It is important to remember that the Work Group provided a forum for program managers to speak out, but it is also true that meetings often focused on review of products such as program descriptions, logic models,

and plausibility analyses, which are seen as "givens" by the program people. The information imbalance limits the value of the discussion; the agenda is set by the evaluator. Input from the manager is needed to insure that the program is accurately portrayed. Sponsor needs tended to focus on standardization of the products. In order to meet sponsor's requirements for uniformity across studies, revisions of products sometimes continued long after a program manager had approved a set of objectives or a model.

3. Formulating Inflexible Procedures

The interest of the sponsor in standardizing the form of the products and procedures in conducting the studies seemed at times to be more important than the substance of the activities. Such attention to form is consistent with the sponsor's goal of demonstrating the method to ED administrators and to program managers as a valuable service, thereby avoiding discontinuation of the service during budget cuts. AIR cooperated in this attempt to prove the value of the service through production of easily recognizable products. The logic models, for example, were viewed as symbols of the trade, and deviations from the prescribed formula were not seriously explored.

Options for streamlining procedures or collapsing steps were seldom considered by AIR or the sponsor, even when the value of a procedure was unclear in meeting a study's objectives. The tendency was to follow a prescribed set of activities because "that's the way you do an E.A." For example, plausibility analyses that laboriously account for each block in the logic model may not always be necessary or useful. This is particularly true if the management options were developed prior to the plausibility analysis, or if the analysis contributes little to other findings. In one case, the staff viewed the analysis as straightforward because the program objectives were agreed upon as plausible, and common sense guided the discussion. At the very minimum, many of the elements of the model could be discussed together. Often, the end result was an analysis that

contained information already presented in another form; the analysis itself was deemed of limited value to the program manager.

One of the chief architects of evaluability assessments states that it is an evolving concept; at least it was in 1979.[4] I think he would agree that rigidity could be its death knell and that informed and responsible flexibility could be a hallmark of its power. Other designers of the system, in their manual for using evaluability assessments, comment on managing such studies:

a. Although a logical and orderly sequence of activities is suggested, it should not be viewed as the only proper one. It is sometimes made possible by the reality of events . . .

b. Steps or analyses defined separately may be done together, and more informally than implied. Even using their own method of analyzing methods prior to evaluations, no two completed assessments looked alike.[5]

Most importantly, they expect the rules to be imperfect and look to the users for necessary changes. Within time and budget constraints, resources and products should be adjusted to serve program managers best.

Cronbach summarizes nicely. "Merit lies not in form of inquiry, but in relevance of information."[6]

4. Using Formal and Informal Knowledge

A major feature of evaluability assessments is determining the extent to which consensus exists on program objectives, major program activities, or appropriate measures of performance. In the early stages of a study, evaluators interview individuals who have different perspectives on the program; the analysis of those findings provides important information about where agreement and disagreement exists on the objectives. This analysis often leads to surprises, but the mechanism for obtaining and using the information is in place.

Knowledge is also obtained informally. Information collected on site consists of both formal "data" and such informal classes as anecdotes and impressions. Sometimes the informal information is of greater value than

the formal knowledge. The point is that the design framework may not accommodate the informal knowledge, a situation not unique to evaluability assessments. The unanticipated information, although interesting, is lost by the formal system and subsequently forgotten.

Other opportunities were lost with site data; some were recognized only after the site reports were written. When a program manager reads information in the site reports that he perceives to be valuable, but cannot see the information used in the management or evaluation options, something is wrong. We seemed to find the information important enough to include in the site report, but perhaps not in the analysis. This is an evaluation failure.

Another neglected area is in the use of the site report as a document. In the current contract, we were required to prepare reports of the site visits. This time consuming and costly activity often continued after the data had been used, for example, to generate management options. What happened to the reports? Copies were sent to local project directors in appreciation of the time they devoted to our visit--or they may eventually be forwarded to the program manager, but typically not early enough to receive timely input after a review of the documents.

These issues are presented to stimulate thought about the application of evaluability assessment methodology to social and educational programs. All of us who have worked on these projects are excited about the process and the interaction it stimulates between program and research communities. We will continue in our attempt to make practice equal performance.

References

1 Cronbach and Associates. Toward reform of program evaluation San Francisco: Jossey-Bass, 1980, p.2.

2 Cronbach, op. cit., p. 11.

3 Wholey, Joseph S. Evaluation: promise and performance. Washington, D.C.: the Urban Institute, 1979, p. 18.

4 Wholey, op. cit., p. 18.

5 Schmidt, R. E., Scanlon, J. W., Bell, J. B. Evaluability assessment: making public programs work better. DHEW Publication OS76-130: Human Services Monograph Series No. 14, 1979, p. 4.

6 Cronbach, op. cit., p. 7.